UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/844,730	04/27/2001	Bin Zhang	10001360-1	1562

7590   10/20/2004

HEWLETT-PACKARD COMPANY
Intellectual Property Administration
P. O. Box 272400
Fort Collins, CO 80527-2400

EXAMINER
COBY, FRANTZ

ART UNIT	PAPER NUMBER
2161	

DATE MAILED: 10/20/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

	Application No.	Applicant(s)
Office Action Summary	09/844,730	ZHANG ET AL.
	Examiner	Art Unit	
	Frantz Coby	2161	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE <u>3</u> MONTH(S) FROM
THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *27 April 2001*.

2a)☐ This action is **FINAL.** 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-20* is/are pending in the application.

    4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-20* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *27 April 2001* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

    Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

    Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

    a)☐ All b)☐ Some * c)☐ None of:

      1.☐ Certified copies of the priority documents have been received.

      2.☐ Certified copies of the priority documents have been received in Application No. _____.

      3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

    * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1)☒ Notice of References Cited (PTO-892)

2)☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3)☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
Paper No(s)/Mail Date <u>04-27-01</u>.

4)☐ Interview Summary (PTO-413)
Paper No(s)/Mail.Date. _____ .

5)☐ Notice of Informal Patent Application (PTO-152)

6)☐ Other: _____.

U.S. Patent and Trademark Office

PTOL-326 (Rev. 1-04) Office Action Summary Part of Paper No./Mail Date 20041016

This is in response to application filed on April 27, 2001 in which claims 1-20 are

presented for examination.

Status of Claims

Claims 1-20 are pending.

Information Disclosure Statement

The information disclosure statement filed on April 27, 2001 is in compliance with

the provisions of 37 CFR 1.97, 1.98 and MPEP § 609. It has been placed in the

application file and the information referred to therein has been considered as to the

merits.

Claim Rejections - 35 USC § 112

The following is a quotation of the second paragraph of 35 U.S.C. 112:

The specification shall conclude with one or more claims particularly pointing out and distinctly
claiming the subject matter which the applicant regards as his invention.

Claims 1-10, and 18 are rejected under 35 U.S.C. 112, second paragraph, as

being indefinite for failing to particularly point out and distinctly claim the subject matter

which applicant regards as the invention.

Regarding claim 1, the preamble renders the claim indefinite because the

preamble recites "A method for clustering data in a system having an integrator ....".

However, the body of the claim is silent on the required steps to arrive with a method for

clustering data using an integrator. The steps recited in the body of the claim can only result in a loop of loading parameters in computers and generate statistics for the purpose of updating the parameters loaded in the computers. Therefore, the steps for clustering data using an integrator needs to be recited in the body of the claim in order to render the present claim 1 definite.

Claims 2-10 are rejected at least for their dependencies directly or indirectly on the rejected claim 1 above.

Claim 3 recites the limitation "its location" in line 4. This claimed feature renders claim 3 indefinite since pronouns are not permitted only what is being referred to by "its" should set forth in the claim.

Claim 4 recites the limitation "the convergence" in lines 2, 3, 5, 7. There is insufficient antecedent basis for this limitation in the claim.

Claim 10 recites the limitation "the data points" in line 1. There is insufficient antecedent basis for this limitation in the claim.

Claim 18 recites the limitation "the convergence" in lines 2, 3, 5, 7. There is insufficient antecedent basis for this limitation in the claim.

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all

obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

This application currently names joint inventors. In considering patentability of

the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of

the various claims was commonly owned at the time any inventions covered therein

were made absent any evidence to the contrary. Applicant is advised of the obligation

under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was

not commonly owned at the time a later invention was made in order for the examiner to

consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g)

prior art under 35 U.S.C. 103(a).

Claims 1-7, 9, 10-12 and 14-20 are rejected under 35 U.S.C. 103(a) as being

unpatentable over Dhillon et al. U.S. Patent no. 6,6,269,376 in view of Fayyad et al.

U.S. Patent no. 6,263,337.

As per claim 1, Dhillon et al. disclose "a method for clustering data in a system

having an integrator and at least two computing units" by providing a method and

system for clustering data in parallel in a distributed memory multiprocessing system

through which data points are clustered in parallel (See Dhillon et al. Title; Abstract). In

particular, Dhillon et al. disclose the claimed feature of "loading each computing unit

with common global parameter values and a particular local data set" (See Dhillon et al.

Col. 7, lines 8-9).

It is noted, however, Dhillon et al. did not specifically detail the claimed features

of "each computing unit generating local sufficient statistics based on the local data set

and global parameter values; and employing the local sufficient statistics of all the

computing units to update the global parameter values" as recited on the instant claim.

On the other hand, Fayyad et al. achieved the aforementioned claimed limitations by

providing a scalable system for expectation maximization clustering of large databases

through which Data contained in the data buffer is used to update the original model

data distributions in each of the K clusters over all M models. Some of the data

belonging to a cluster is summarized or compressed and stored as a reduced form of

the data representing sufficient statistics of the data. More data is accessed from the

database and the models are updated. An updated set of parameters for the clusters is

determined from the summarized data (sufficient statistics) and the newly acquired data

(See Fayyad et al. Abstract; Col. 3, lines 10-24).

It would have been obvious to one of ordinary skill in the art at the time of the

invention to modify the clustering system of Dhillon et al. by incorporating the

methodology of updating parameters based on statistics generated by the processes as

taught by Fayyad (See Fayyad et al. Abstract) into the k-means process provided

thereof (See Dhillon et al. Columns 5-6).

The motivation being to permit the clustering system of Dhillon et al. to provide a

process of gathering data from databases and characterizing the data clusters based on

newly sampled data from the database (See Fayyad et al. Col. 3, lines 30-34).

As per claim 2, most of the limitations of this claim have been noted in the

rejection of claim 1. Applicant's attention is directed to the rejection of claim 1 above.

In addition, Dhillon et al. disclose the claimed features of "loading each computing unit

with common global parameter values comprises: and a particular local data set further

receiving a set of data points to be clustered; dividing the data points into at least two

local data sets; sending common global parameter values to each of the computing

units', and sending each local data sets to a designated computing unit" (See Dhillon et

al. Abstract; Col 1, line 54-Col. 2, line 18).

As per claim 3, most of the limitations of this claim have been noted in the

rejection of claim 2. Applicant's attention is directed to the rejection of claim 2 above.

In addition, Dhillon et al. disclose the claimed features of "each computing unit

integrator sending its local sufficient statistics to the integrator; the integrator

determining global sufficient statistics based on the local sufficient statistics of all the

computing units; and the integrator determining updated global parameter values based

on the global sufficient statistics" (See Fayyad et al. Figure 1) which is a computer

system for use in practicing the invention. This is therefore, a clear indication that each

computer in the network is able to send its local sufficient statistics. Also, the Applicant

should duly note that Figure 1 shows a network of computers and the text of Fayyad et al. specifically stated that an updated set of parameters for the clusters is determined from the summarized data sufficient statistics (See Fayyad et al. Abstract). Therefore, the aspect of determining updated global parameter values based on the global sufficient statistics is primarily incorporated in the clustering system of Fayyad et al.

As per claim 4, most of the limitations of this claim have been noted in the rejection of claim 1. Applicant's attention is directed to the rejection of claim 1 above. In addition, Dhillon et al. disclose the claimed features of "checking the convergence quality; determining whether the convergence meets a predetermined quality; and when the convergence meets a predetermined quality, stop processing; otherwise; when the convergence fails to meet a predetermined quality, providing the updated global parameter values to the computing units and repeating steps (a) to (c)" (See Dhillon et al. Col. 5, line 20-Col. 5, line 44).

As per claims 5-6, most of the limitations of these claims have been noted in the rejection of claim 2. Applicant's attention is directed to the rejection of claim 2 above. In addition, Dhillon et al. disclose the claimed features of "wherein sending common global parameter values to each of the computing units includes the step of: broadcasting common global parameter values to each of the computing units;

" initializing the common global parameter values before sending the common global

parameter values to each of the computing units" (See Dhillon et al. Col. 6, line 50-Col.

7, line 18).

As per claims 7 and 10, most of the limitations of these claims have been noted

in the rejection of claim 1. Applicant's attention is directed to the rejection of claim 1

above. In addition, Dhillon et al. disclose the claimed features of "wherein a distributed

K-Means clustering algorithm is implemented"; "wherein the data points to be clustered

are naturally distributed" (See Dhillon et al. Figure 2; Col. 5, line 13-Col. 6, line 49).

As per claim 9, most of the limitations of this claim have been noted in the

rejection of claim 1. Applicant's attention is directed to the rejection of claim 1 above.

In addition, Fayyad et al. disclose the claimed features of "wherein a distributed

Expectation-Maximization (EM) clustering algorithm is implemented" (See Fayyad et al.

Col. 5, line 35-Col. 18, line 46).

As per claim 11, all the limitations of this claim have been noted in the rejection

of claim 1. It is therefore rejected as set forth above.

As per claim 12, all the limitations of this claim have been noted in the rejection

of claim 7. It is therefore rejected as set forth above.

As per claim 14, all the limitations of this claim have been noted in the rejection of claim 9. It is therefore rejected as set forth above.

As per claim 15, all the limitations of this claim have been noted in the rejection of claims 7 and 10. It is therefore rejected as set forth above.

As per claim 16, all the limitations of this claim have been noted in the rejection of claim 2. It is therefore rejected as set forth above.

As per claim 17, all the limitations of this claim have been noted in the rejection of claim 3. It is therefore rejected as set forth above.

As per claim 18, all the limitations of this claim have been noted in the rejection of claim 4. It is therefore rejected as set forth above.

As per claim 19, all the limitations of this claim have been noted in the rejection of claims 5-6. It is therefore rejected as set forth above.

As per claim 20, all the limitations of this claim have been noted in the rejection of claims 5-6. It is therefore rejected as set forth above.

Claims 8 and 13 are rejected under 35 U.S.C. 103(a) as being unpatentable over

Dhillon et al. U.S. Patent no. 6,6,269,376 in view of Fayyad et al. U.S. Patent no.

6,263,337 as applied to claims 1-7 and 9-10 above and further in view of Zhang et al.

U.S. Patent no. 6,584,433.

As per claim 8, most of the limitations of this claim have been noted in the

rejection of claim 1. Applicant's attention is directed to the rejection of claim 1 above.

It is noted, however, neither Dhillon et al. nor Fayyad et al. specifically detail the

aspect of "wherein a distributed K-Harmonic Means clustering algorithm is

implemented" as recited in the instant claim. On the other hand, Zhang et al. disclosed

a harmonic based clustering method and system that implements that recognizes that

that K-Means and Expectation Maximization (EM) are two prior art methods for data

clustering (See Zhang et al. Col. 1, lines 65-67) and implements a K-Harmonic Means

performance function (See Zhang et al. Col. 3, lines 11-13).

It would have been obvious to one of ordinary skill in the art at the time of the

invention to improved on improve on the clustering systems of Dhillon et al. and Fayyad

et al. by implementing a K-harmonic because such modification would provide a

clustering method and system for reducing the dependency of clustering results to the

initialization centers; thus, improving the quality of clustering results or the convergence

quality of the clustering , the convergence rate of the clustering (See Zhang et al. Col. 2,

lines 55-65).

As per claim 13, all the limitations of this claim have been noted in the rejection

of claim 8. It is therefore rejected as set forth above.

Conclusion

Any inquiry concerning this communication or earlier communications from the

examiner should be directed to Frantz Coby whose telephone number is (571) 272-

4017. The examiner can normally be reached on Monday-Friday 3:00 P.M. - 11:00

P.M..

If attempts to reach the examiner by telephone are unsuccessful, the examiner's

supervisor, Safet Metjahic can be reached on 703 308 1436. The fax phone number for

the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the

Patent Application Information Retrieval (PAIR) system. Status information for

published applications may be obtained from either Private PAIR or Public PAIR.

Status information for unpublished applications is available through Private PAIR only.

For more information about the PAIR system, see http://pair-direct.uspto.gov. Should

you have questions on access to the Private PAIR system, contact the Electronic

Business Center (EBC) at 866-217-9197 (toll-free).

Frantz Coby
Primary Examiner
Art Unit 2171

October 16, 2004